Supplementary Materials

Appendix A.1. Proof of the Proposition 1.

Proof:

For simplification, we define $\widetilde{\boldsymbol{\mathcal{F}}} = \left[\!\left[\boldsymbol{\mathcal{F}}; \boldsymbol{A}_i^{(1)}, \boldsymbol{A}_i^{(2)}, \boldsymbol{A}_i^{(3)}\right]\!\right]$, $\widetilde{\boldsymbol{\mathcal{R}}}_i = \left[\!\left[\boldsymbol{\mathcal{R}}_i; \boldsymbol{B}_i^{(1)}, \boldsymbol{B}_i^{(2)}, \boldsymbol{B}_i^{(3)}\right]\!\right]$

$$\boldsymbol{\mathcal{R}}_i \backsim \boldsymbol{N}_{P_2,Q_2,R_2}(\boldsymbol{\mathcal{O}}; \boldsymbol{\Sigma}_{\mathrm{r}}, \boldsymbol{\Psi}_{\mathrm{r}}, \boldsymbol{\Omega}_{\mathrm{r}}) \text{ if } \mathrm{vec}(\boldsymbol{\mathcal{R}}_i) \backsim \boldsymbol{N}_{P_2Q_2R_2}(\mathrm{vec}(\boldsymbol{\mathcal{O}}), \boldsymbol{\Omega}_{\mathrm{r}} \otimes \boldsymbol{\Psi}_{\mathrm{r}} \otimes \boldsymbol{\Sigma}_{\mathrm{r}})$$

$$\boldsymbol{\mathcal{Y}}_i = \boldsymbol{\mathcal{F}} \times_1 \boldsymbol{A}_i^{(1)} \times_2 \boldsymbol{A}_i^{(2)} \times_3 \boldsymbol{A}_i^{(3)} + \boldsymbol{\mathcal{R}}_i \times_1 \boldsymbol{B}_i^{(1)} \times_2 \boldsymbol{B}_i^{(2)} \times_3 \boldsymbol{B}_i^{(3)} + \boldsymbol{\mathcal{E}}_i$$

We check the distribution of random effects,

$$\mathrm{vec}\left(\boldsymbol{\mathcal{R}}_i \times_1 \boldsymbol{B}_i^{(1)} \times_2 \boldsymbol{B}_i^{(2)} \times_3 \boldsymbol{B}_i^{(3)}\right) = \left(\boldsymbol{B}_i^{(3)} \otimes \boldsymbol{B}_i^{(2)} \otimes \boldsymbol{B}_i^{(1)}\right)\mathrm{vec}(\boldsymbol{\mathcal{R}}_i)$$

$$\backsim \boldsymbol{N}_{IJK}\left(\left(\boldsymbol{B}_i^{(3)} \otimes \boldsymbol{B}_i^{(2)} \otimes \boldsymbol{B}_i^{(1)}\right)\mathrm{vec}(\boldsymbol{\mathcal{O}}), \left(\boldsymbol{B}_i^{(3)} \otimes \boldsymbol{B}_i^{(2)} \otimes \boldsymbol{B}_i^{(1)}\right)(\boldsymbol{\Omega}_{\mathrm{r}} \otimes \boldsymbol{\Psi}_{\mathrm{r}} \otimes \boldsymbol{\Sigma}_{\mathrm{r}})\left(\boldsymbol{B}_i^{(3)} \otimes \boldsymbol{B}_i^{(2)} \otimes \boldsymbol{B}_i^{(1)}\right)^T\right)$$

$$\backsim \boldsymbol{N}_{IJK}\left(\mathrm{vec}(\boldsymbol{\mathcal{O}}), \left(\boldsymbol{B}_i^{(3)} \boldsymbol{\Omega}_{\mathrm{r}} \boldsymbol{B}_i^{(3)^T}\right) \otimes \left(\boldsymbol{B}_i^{(2)} \boldsymbol{\Psi}_{\mathrm{r}} \boldsymbol{B}_i^{(2)^T}\right) \otimes \left(\boldsymbol{B}_i^{(1)} \boldsymbol{\Sigma}_{\mathrm{r}} \boldsymbol{B}_i^{(1)^T}\right)\right).$$

Then $\widetilde{\boldsymbol{\mathcal{R}}}_i = \boldsymbol{\mathcal{R}}_i \times_1 \boldsymbol{B}_i^{(1)} \times_2 \boldsymbol{B}_i^{(2)} \times_3 \boldsymbol{B}_i^{(3)}$ follows tensor normal distribution

$$\widetilde{\boldsymbol{\mathcal{R}}}_i \backsim \boldsymbol{N}_{J,K,L}\left(\boldsymbol{\mathcal{O}}; \boldsymbol{B}_i^{(1)} \boldsymbol{\Sigma}_{\mathrm{r}} \boldsymbol{B}_i^{(1)^T}, \boldsymbol{B}_i^{(2)} \boldsymbol{\Psi}_{\mathrm{r}} \boldsymbol{B}_i^{(2)^T}, \boldsymbol{B}_i^{(3)} \boldsymbol{\Omega}_{\mathrm{r}} \boldsymbol{B}_i^{(3)^T}\right)$$

Since the random effects core tensor and residual errors tensor are independent of each other, $\boldsymbol{\mathcal{E}}_i \backsim \boldsymbol{N}_{J,K,L}(\boldsymbol{\mathcal{O}}; \boldsymbol{\Sigma}_{\varepsilon}, \boldsymbol{\Psi}_{\varepsilon}, \boldsymbol{\Omega}_{\varepsilon})$, the three dimensional joint tensor $\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{R}}}_i & \boldsymbol{\mathcal{O}} \\ \boldsymbol{\mathcal{O}} & \boldsymbol{\mathcal{E}}_i \end{bmatrix}$ satisfies

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{R}}}_i & \boldsymbol{\mathcal{O}} \\ \boldsymbol{\mathcal{O}} & \boldsymbol{\mathcal{E}}_i \end{bmatrix} \backsim \boldsymbol{N}_{2J,2K,2L}\left(\boldsymbol{\mathcal{O}}; \begin{bmatrix} \boldsymbol{B}_i^{(1)} \boldsymbol{\Sigma}_{\mathrm{r}} \boldsymbol{B}_i^{(1)^T} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\varepsilon} \end{bmatrix}, \begin{bmatrix} \boldsymbol{B}_i^{(2)} \boldsymbol{\Sigma}_{\mathrm{r}} \boldsymbol{B}_i^{(2)^T} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{\varepsilon} \end{bmatrix}, \begin{bmatrix} \boldsymbol{B}_i^{(3)} \boldsymbol{\Sigma}_{\mathrm{r}} \boldsymbol{B}_i^{(3)^T} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{\varepsilon} \end{bmatrix}\right)$$

Based on the Theorem 3.1 (Ohlson et al. 2011), we define $\mathcal{A} = [\boldsymbol{I}_J \quad \boldsymbol{I}_J] \otimes [\boldsymbol{I}_K \quad \boldsymbol{I}_K] \otimes [\boldsymbol{I}_L \quad \boldsymbol{I}_L] \in \boldsymbol{\mathcal{T}}_{\otimes}^{[J,K,L],[2J,2K,2L]}$, then

$$\mathcal{A}\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{R}}}_i & \boldsymbol{\mathcal{O}} \\ \boldsymbol{\mathcal{O}} & \boldsymbol{\mathcal{E}}_i \end{bmatrix} = \widetilde{\boldsymbol{\mathcal{R}}}_i + \boldsymbol{\mathcal{E}}_i \backsim \boldsymbol{N}_{J,K,L}\left(\boldsymbol{\mathcal{O}}; \boldsymbol{B}_i^{(1)} \boldsymbol{\Sigma}_{\mathrm{r}} \boldsymbol{B}_i^{(1)^T} + \boldsymbol{\Sigma}_{\varepsilon}, \boldsymbol{B}_i^{(2)} \boldsymbol{\Psi}_{\mathrm{r}} \boldsymbol{B}_i^{(2)^T} + \boldsymbol{\Psi}_{\varepsilon}, \boldsymbol{B}_i^{(3)} \boldsymbol{\Omega}_{\mathrm{r}} \boldsymbol{B}_i^{(3)^T} + \boldsymbol{\Omega}_{\varepsilon}\right)$$

Thus

$$\boldsymbol{\mathcal{Y}}_i \backsim \boldsymbol{N}_{J,K,L}\left(\left[\!\left[\boldsymbol{\mathcal{F}}; \boldsymbol{A}_i^{(1)}, \boldsymbol{A}_i^{(2)}, \boldsymbol{A}_i^{(3)}\right]\!\right]; \boldsymbol{B}_i^{(1)} \boldsymbol{\Sigma}_{\mathrm{r}} \boldsymbol{B}_i^{(1)^T} + \boldsymbol{\Sigma}_{\varepsilon}, \boldsymbol{B}_i^{(2)} \boldsymbol{\Psi}_{\mathrm{r}} \boldsymbol{B}_i^{(2)^T} + \boldsymbol{\Psi}_{\varepsilon}, \boldsymbol{B}_i^{(3)} \boldsymbol{\Omega}_{\mathrm{r}} \boldsymbol{B}_i^{(3)^T} + \boldsymbol{\Omega}_{\varepsilon}\right).$$

Appendix A.2. Proof of the Proposition 2.

Proof:

The likelihood function for Equation (3) is shown as following:

$L_1 = (2\pi)^{-\frac{JKL}{2}} \cdot |\boldsymbol{\Omega}_i \otimes \boldsymbol{\Psi}_i|^{-\frac{J}{2}} \cdot |\boldsymbol{\Sigma}_i|^{-\frac{KL}{2}} \cdot \exp\left(-\frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Sigma}_i^{-1}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right)^T (\boldsymbol{\Omega}_i \otimes \boldsymbol{\Psi}_i)^{-1}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right)\right]\right) = (2\pi)^{-\frac{JKL}{2}} \cdot |\boldsymbol{\Omega}_i|^{-\frac{JK}{2}} \cdot |\boldsymbol{\Psi}_i|^{-\frac{JL}{2}} \cdot |\boldsymbol{\Sigma}_i|^{-\frac{KL}{2}} \cdot \exp\left(-\frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Sigma}_i^{-1}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right)^T (\boldsymbol{\Omega}_i \otimes \boldsymbol{\Psi}_i)^{-1}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right)\right]\right).$

Similarly, we can get the likelihood function for Equations (4) and (5):

$L_2 = (2\pi)^{-\frac{JKL}{2}} \cdot |\boldsymbol{\Omega}_i|^{-\frac{JK}{2}} \cdot |\boldsymbol{\Psi}_i|^{-\frac{JL}{2}} \cdot |\boldsymbol{\Sigma}_i|^{-\frac{KL}{2}} \cdot \exp\left(-\frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Psi}_i^{-1}\left(\boldsymbol{Y}_{i(2)} - \boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(1)}\right)^T\right)^T (\boldsymbol{\Omega}_i \otimes \boldsymbol{\Sigma}_i)^{-1}\left(\boldsymbol{Y}_{i(2)} - \boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(1)}\right)^T\right)\right]\right);$

$L_3 = (2\pi)^{-\frac{JKL}{2}} \cdot |\boldsymbol{\Omega}_i|^{-\frac{JK}{2}} \cdot |\boldsymbol{\Psi}_i|^{-\frac{JL}{2}} \cdot |\boldsymbol{\Sigma}_i|^{-\frac{KL}{2}} \cdot \exp\left(-\frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Omega}_i^{-1}\left(\boldsymbol{Y}_{i(3)} - \boldsymbol{A}_i^{(3)}\boldsymbol{F}_{(3)}\left(\boldsymbol{A}_i^{(2)} \otimes \boldsymbol{A}_i^{(1)}\right)^T\right)^T (\boldsymbol{\Psi}_i \otimes \boldsymbol{\Sigma}_i)^{-1}\left(\boldsymbol{Y}_{i(3)} - \boldsymbol{A}_i^{(3)}\boldsymbol{F}_{(3)}\left(\boldsymbol{A}_i^{(2)} \otimes \boldsymbol{A}_i^{(1)}\right)^T\right)\right]\right).$

To prove the log-likelihood functions of Equations (3-5) are same, we need to show the parts within th[$\cdot$] are same.

Considering the commutation matrix $\boldsymbol{K}_{L,JK}$ and $\boldsymbol{K}_{R_1,P_1Q_1}$, we have

$$\mathrm{vec}\left(\boldsymbol{Y}_{i(1)}\right) = \boldsymbol{K}_{L,JK}\mathrm{vec}\left(\boldsymbol{Y}_{i(3)}\right)$$

$$\mathrm{vec}\left(\boldsymbol{F}_{(1)}\right) = \boldsymbol{K}_{R_1,P_1Q_1}\mathrm{vec}\left(\boldsymbol{F}_{(3)}\right)$$

$$\mathrm{vec}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right) = \boldsymbol{K}_{K,JL}\mathrm{vec}\left(\boldsymbol{Y}_{i(2)} - \boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(1)}\right)^T\right)$$

$\mathrm{tr}\left[\boldsymbol{\Sigma}_i^{-1}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right)^T (\boldsymbol{\Omega}_i \otimes \boldsymbol{\Psi}_i)^{-1}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right)\right] = \mathrm{vec}^T\left[\left(\boldsymbol{\Sigma}_i^{-1}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right)^T (\boldsymbol{\Omega}_i \otimes \boldsymbol{\Psi}_i)^{-1}\right)^T\right]\mathrm{vec}\left[\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right] = \mathrm{vec}^T\left[\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right] \cdot (\boldsymbol{\Omega}_i^{-1} \otimes \boldsymbol{\Psi}_i^{-1} \otimes \boldsymbol{\Sigma}_i^{-1})^T \cdot \mathrm{vec}\left[\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(2)}\right)^T\right] = \mathrm{vec}^T\left[\boldsymbol{Y}_{i(2)} - \boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)} \otimes \boldsymbol{A}_i^{(1)}\right)^T\right]\boldsymbol{K}_{K,JL}^T \cdot$

$(\boldsymbol{\Omega}_i^{-1}\otimes\boldsymbol{\Psi}_i^{-1}\otimes\boldsymbol{\Sigma}_i^{-1})^T\cdot\boldsymbol{K}_{K,JL}\mathrm{vec}\left[\boldsymbol{Y}_{i(2)}-\boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right]=\mathrm{vec}^T\left[\boldsymbol{Y}_{i(2)}-\right.$

$\left.\boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right]\cdot(\boldsymbol{\Omega}_i^{-1}\otimes\boldsymbol{\Sigma}_i^{-1}\otimes\boldsymbol{\Psi}_i^{-1})^T\cdot\mathrm{vec}\left[\boldsymbol{Y}_{i(2)}-\boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right]=$

$\mathrm{vec}^T\left[\left(\boldsymbol{\Psi}_i^{-1}\left(\boldsymbol{Y}_{i(2)}-\boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right)^T(\boldsymbol{\Omega}_i\otimes\boldsymbol{\Sigma}_i)^{-1}\right)^T\right]\mathrm{vec}\left[\boldsymbol{Y}_{i(2)}-\right.$

$\left.\boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right]=\mathrm{tr}\left[\boldsymbol{\Psi}_i^{-1}\left(\boldsymbol{Y}_{i(2)}-\boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right)^T(\boldsymbol{\Omega}_i\otimes\boldsymbol{\Sigma}_i)^{-1}\left(\boldsymbol{Y}_{i(2)}-\right.\right.$

$\left.\left.\boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right)\right].$

Thus, $L_1=L_2$. Similarly, we can prove $L_3=L_2$.

According to the result above, we can get the log-likelihood function as

$\boldsymbol{l}_i=-\frac{JKL}{2}\log 2\pi-\frac{JK}{2}\log|\boldsymbol{\Omega}_i|-\frac{JL}{2}\log|\boldsymbol{\Psi}_i|-\frac{KL}{2}\log|\boldsymbol{\Sigma}_i|-\frac{1}{2}\left(\mathrm{vec}\left(\boldsymbol{Y}_{i(1)}-\right.\right.$

$\left.\left.\boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(2)}\right)^T\right)\right)^T(\boldsymbol{\Omega}_i^{-1}\otimes\boldsymbol{\Psi}_i^{-1}\otimes\boldsymbol{\Sigma}_i^{-1})\mathrm{vec}\left(\boldsymbol{Y}_{i(1)}-\boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(2)}\right)^T\right).$

Appendix A.3. Proof of the Proposition 3.

Proof:

Given the response tensors $\boldsymbol{\mathcal{Y}}_i$, the basis $\boldsymbol{A}_i^{(1)}$, $\boldsymbol{A}_i^{(2)}$, $\boldsymbol{A}_i^{(3)}$ with $i=1,\cdots,N$, the log-likelihood function for all the samples is

$\boldsymbol{l}=\sum_{i=1}^N\boldsymbol{l}_i=\sum_{i=1}^N\left\{-\frac{JKL}{2}\log 2\pi-\frac{JK}{2}\log|\boldsymbol{\Omega}_i|-\frac{JL}{2}\log|\boldsymbol{\Psi}_i|-\frac{KL}{2}\log|\boldsymbol{\Sigma}_i|-\frac{1}{2}\left(\mathrm{vec}\left(\boldsymbol{Y}_{i(1)}-\right.\right.\right.$

$\left.\left.\left.\boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(2)}\right)^T\right)\right)^T(\boldsymbol{\Omega}_i^{-1}\otimes\boldsymbol{\Psi}_i^{-1}\otimes\boldsymbol{\Sigma}_i^{-1})\mathrm{vec}\left(\boldsymbol{Y}_{i(1)}-\boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(2)}\right)^T\right)\right\}.$

We take the first derivative of the log-likelihood function with respect to $\mathrm{vec}(\widehat{\boldsymbol{\mathcal{F}}})$ is

$\frac{\boldsymbol{dl}}{\boldsymbol{d}\mathrm{vec}(\widehat{\boldsymbol{\mathcal{F}}})}=\sum_{i=1}^N\mathrm{vec}\left[\boldsymbol{\mathcal{Y}}_i-\widehat{\widehat{\boldsymbol{\mathcal{F}}}}\right]\cdot(\boldsymbol{\Omega}_i^{-1}\otimes\boldsymbol{\Psi}_i^{-1}\otimes\boldsymbol{\Sigma}_i^{-1})\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(2)}\otimes\boldsymbol{A}_i^{(1)}\right).$

Let the first derivative above equals to zero, we can get the maximum likelihood estimator of $\mathrm{vec}(\boldsymbol{\mathcal{F}})$ is

$$\mathrm{vec}(\widehat{\boldsymbol{\mathcal{F}}}) = \left(\sum_{i=1}^{N}\left(\boldsymbol{A}_i^{(3)^T}\boldsymbol{\Omega}_i^{-1}\boldsymbol{A}_i^{(3)}\right)\otimes\left(\boldsymbol{A}_i^{(2)^T}\boldsymbol{\Psi}_i^{-1}\boldsymbol{A}_i^{(2)}\right)\otimes\left(\boldsymbol{A}_i^{(1)^T}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{A}_i^{(1)}\right)\right)^{-1}$$
$$\cdot\left(\sum_{i=1}^{N}\left(\boldsymbol{A}_i^{(3)^T}\boldsymbol{\Omega}_i^{-1}\right)\otimes\left(\boldsymbol{A}_i^{(2)^T}\boldsymbol{\Psi}_i^{-1}\right)\otimes\left(\boldsymbol{A}_i^{(1)^T}\boldsymbol{\Sigma}_i^{-1}\right)\cdot\mathrm{vec}(\boldsymbol{\mathcal{Y}}_i)\right)$$

Moreover, we can show that the estimator $\mathrm{vec}(\widehat{\boldsymbol{\mathcal{F}}})$ given in Equation (7) is uniquely determined regardless of the parametrization of the covariance matrices. Because when $\boldsymbol{\Psi}_i, i = 1,\cdots,N$ is replaced with $\mathrm{m}\boldsymbol{\Psi}_i$ and $\mathrm{m} \in \mathbb{R}^+$, the expression $\mathrm{vec}(\widehat{\boldsymbol{\mathcal{F}}})$ above still satisfies.

Assume that $\boldsymbol{B}_i^{(1)}, \boldsymbol{B}_i^{(2)}, \boldsymbol{B}_i^{(3)}$ are constant for all $i = 1,\cdots,N$, and define that $\boldsymbol{B}_i^{(1)} = \boldsymbol{B}^{(1)}, \boldsymbol{B}_i^{(2)} = \boldsymbol{B}^{(2)}, \boldsymbol{B}_i^{(3)} = \boldsymbol{B}^{(3)}$ for $i = 1,\cdots,N$, We take the first derivatives of the log-likelihood functions with respect to $\boldsymbol{\Sigma}_i$, $\boldsymbol{\Psi}_i$, $\boldsymbol{\Omega}_i$ are

$$\frac{\boldsymbol{dl}}{\boldsymbol{d\Sigma}_i} = \frac{KLN}{2}\boldsymbol{\Sigma}_{\boldsymbol{i}}^{-\mathbf{1}} - \boldsymbol{\Sigma}_{\boldsymbol{i}}^{-\mathbf{1}}\left\{\sum_{i=1}^{N}\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(2)}\right)^T\right)(\boldsymbol{\Omega}_i^{-1}\otimes\boldsymbol{\Psi}_i^{-1})\left(\boldsymbol{Y}_{i(1)} - \boldsymbol{A}_i^{(1)}\boldsymbol{F}_{(1)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(2)}\right)^T\right)^T\right\}\boldsymbol{\Sigma}_{\boldsymbol{i}}^{-\mathbf{1}}.$$

$$\frac{\boldsymbol{dl}}{\boldsymbol{d\Psi}_i} = \frac{JLN}{2}\boldsymbol{\Psi}_{\boldsymbol{i}}^{-\mathbf{1}} - \boldsymbol{\Psi}_{\boldsymbol{i}}^{-\mathbf{1}}\left\{\sum_{i=1}^{N}\left(\boldsymbol{Y}_{i(2)} - \boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right)(\boldsymbol{\Omega}_i^{-1}\otimes\boldsymbol{\Sigma}_i^{-1})\left(\boldsymbol{Y}_{i(2)} - \boldsymbol{A}_i^{(2)}\boldsymbol{F}_{(2)}\left(\boldsymbol{A}_i^{(3)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right)^T\right\}\boldsymbol{\Psi}_{\boldsymbol{i}}^{-\mathbf{1}}.$$

$$\frac{\boldsymbol{dl}}{\boldsymbol{d\Omega}_i} = \frac{KJN}{2}\boldsymbol{\Omega}_{\boldsymbol{i}}^{-\mathbf{1}} - \boldsymbol{\Omega}_{\boldsymbol{i}}^{-\mathbf{1}}\left\{\sum_{i=1}^{N}\left(\boldsymbol{Y}_{i(3)} - \boldsymbol{A}_i^{(3)}\boldsymbol{F}_{(3)}\left(\boldsymbol{A}_i^{(2)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right)(\boldsymbol{\Psi}_i^{-1}\otimes\boldsymbol{\Sigma}_i^{-1})\left(\boldsymbol{Y}_{i(3)} - \boldsymbol{A}_i^{(3)}\boldsymbol{F}_{(3)}\left(\boldsymbol{A}_i^{(2)}\otimes\boldsymbol{A}_i^{(1)}\right)^T\right)^T\right\}\boldsymbol{\Omega}_{\boldsymbol{i}}^{-\mathbf{1}}.$$

Letting the first derivatives of the log-likelihood functions with respect to $\boldsymbol{\Sigma}_i$, $\boldsymbol{\Psi}_i$, $\boldsymbol{\Omega}_i$ be zeros, we can get the maximum likelihood estimators of $\boldsymbol{\Sigma}_i$, $\boldsymbol{\Psi}_i$, $\boldsymbol{\Omega}_i$ are

$$\widehat{\boldsymbol{\Sigma}}_i = \frac{1}{KLN}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{Y}}_i - \widehat{\boldsymbol{\mathcal{F}}}\right)_{(1)}\cdot(\widehat{\boldsymbol{\Omega}}_i^{-1}\otimes\widehat{\boldsymbol{\Psi}}_i^{-1})\cdot\left(\boldsymbol{\mathcal{Y}}_i - \widehat{\boldsymbol{\mathcal{F}}}\right)_{(1)}^T;$$

$$\widehat{\boldsymbol{\Psi}}_i = \frac{1}{JLN}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{Y}}_i - \widehat{\widehat{\boldsymbol{\mathcal{F}}}}\right)_{(2)}\cdot(\widehat{\boldsymbol{\Omega}}_i^{-1}\otimes\widehat{\boldsymbol{\Sigma}}_i^{-1})\cdot\left(\boldsymbol{\mathcal{Y}}_i - \widehat{\widehat{\boldsymbol{\mathcal{F}}}}\right)_{(2)}^T;$$

$$\widehat{\boldsymbol{\Omega}}_i = \frac{1}{JKN}\sum_{i=1}^{N}\left(\boldsymbol{\mathcal{Y}}_i - \widehat{\widehat{\boldsymbol{\mathcal{F}}}}\right)_{(3)}\cdot(\widehat{\boldsymbol{\Psi}}_i^{-1}\otimes\widehat{\boldsymbol{\Sigma}}_i^{-1})\cdot\left(\boldsymbol{\mathcal{Y}}_i - \widehat{\widehat{\boldsymbol{\mathcal{F}}}}\right)_{(3)}^T.$$

Straightforwardly, if both $\boldsymbol{A}_i^{(1)}, \boldsymbol{A}_i^{(2)}, \boldsymbol{A}_i^{(3)}$ and $\boldsymbol{B}_i^{(1)}, \boldsymbol{B}_i^{(2)}, \boldsymbol{B}_i^{(3)}$ are constant for all $i = 1, \cdots, N$, the maximum likelihood estimators of $\boldsymbol{\Sigma}_i, \boldsymbol{\Psi}_i, \boldsymbol{\Omega}_i$ are

$\widehat{\boldsymbol{\Sigma}}_i = \frac{1}{KLN} \sum_{i=1}^{N} (\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{\mathcal{Y}}})_{(1)} \cdot \left(\widehat{\boldsymbol{\Omega}}_i^{-1} \otimes \widehat{\boldsymbol{\Psi}}_i^{-1}\right) \cdot (\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{\mathcal{Y}}})_{(1)}^T;$

$\widehat{\boldsymbol{\Psi}}_i = \frac{1}{JLN} \sum_{i=1}^{N} (\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{\mathcal{Y}}})_{(2)} \cdot \left(\widehat{\boldsymbol{\Omega}}_i^{-1} \otimes \widehat{\boldsymbol{\Sigma}}_i^{-1}\right) \cdot (\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{\mathcal{Y}}})_{(2)}^T;$

$\widehat{\boldsymbol{\Omega}}_i = \frac{1}{JKN} \sum_{i=1}^{N} (\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{\mathcal{Y}}})_{(3)} \cdot \left(\widehat{\boldsymbol{\Psi}}_i^{-1} \otimes \widehat{\boldsymbol{\Sigma}}_i^{-1}\right) \cdot (\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{\mathcal{Y}}})_{(3)}^T.$